Between-Patient and Within-Patient (Site-to-Site) Variability in Estrogen Receptor Binding, Measured In Vivo by ¹⁸F-Fluoroestradiol PET

Brenda F. Kurland¹, Lanell M. Peterson², Jean H. Lee², Hannah M. Linden³, Erin K. Schubert², Lisa K. Dunnwald², Jeanne M. Link², Kenneth A. Krohn², and David A. Mankoff²

¹Department of Clinical Statistics, Fred Hutchinson Cancer Research Center, Seattle Washington; ²Department of Radiology, University of Washington, Seattle, Washington; and ³Division of Medical Oncology, University of Washington, Seattle, Washington

Heterogeneity of estrogen receptor (ER) expression may be an important predictor of breast cancer therapeutic response. ¹⁸F-fluoroestradiol PET produces in vivo quantitative measurements of regional estrogen binding in breast cancer tumors. We describe within-patient (site-to-site) and between-patient heterogeneity of lesions in patients scheduled to receive endocrine therapy. Methods: In 91 patients with a prior ER-positive biopsy, 505 lesions were analyzed for both ¹⁸F-fluoroestradiol and ¹⁸F-FDG uptake and the ¹⁸F-fluoroestradiol/¹⁸F-FDG uptake ratio. Standardized uptake values (SUVs) were recorded for up to 16 lesions per patient, of 1.5 cm or more and visible on ¹⁸F-FDG PET or conventional staging. Linear mixed-effects regression models examined associations between PET parameters and patient or lesion characteristics and estimated variance components. A reader study of SUV measurements for 9 scans further examined sources of within-patient variability. Results: Average ¹⁸F-fluoroestradiol uptake and ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio varied greatly across these patients, despite a history of ER-positive disease: about 37% had low or absent ¹⁸F-fluoroestradiol uptake even with marked ¹⁸F-FDG uptake. ¹⁸F-fluoroestradiol SUV and ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio measurements within patients with multiple lesions were clustered around the patient’s average value in most cases. Summarizing these findings, the intraclass correlation coefficient (proportion of total variation that is between-patient) was 0.60 (95% confidence interval, 0.50–0.69) for ¹⁸F-fluoroestradiol SUV and 0.65 (95% confidence interval, 0.56–0.73) for the ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio. Some within-patient variation in PET measures (22%–44%) was attributable to interobserver variability as measured by the reader study. A subset of patients had mixed uptake, with widely disparate ¹⁸F-fluoroestradiol SUV or ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio for lesions in the same scan. Conclusion: ¹⁸F-fluoroestradiol uptake and the ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio varied greatly between patients but were usually consistent across lesions in the same scan. The average ¹⁸F-fluoroestradiol SUV and ¹⁸F-fluoroestradiol/¹⁸F-FDG ratio for a limited sample of lesions appear to provide a reasonable summary of synchronous ER expression for most patients. However, imaging the entire disease burden remains important to identify the subset of patients with mixed uptake, who may be at a critical point in their disease evolution.

Key Words: estrogen receptor imaging; ¹⁸F-FES PET; heterogeneity; breast cancer

J Nucl Med 2011; 52:1–9
DOI: 10.2967/jnumed.111.091439

More than 70% of breast cancers are estrogen receptor–positive (ER+), and ER-directed adjuvant therapy is credited as a key factor in declines in breast cancer mortality. However, not all patients respond to endocrine therapy, and most initial responders later become refractory: the objective response rate to second-line endocrine therapy is less than 20% (1). One factor hypothesized to underlie unpredictable response and acquired resistance to endocrine therapy is heterogeneity of ER expression. Because levels of ER expression may be heterogeneous within the same patient (2–4), biopsy of a single tumor site may not be representative. ER expression may also change over time in the same patient, caused either by genetic or epigenetic loss of the receptor expression as the tumor grows and metastasizes or as the selective pressure of endocrine therapy allows the emergence of clones with low or absent ER expression (5,6). Alternatively, ER expression may be stable, but other proliferative or antiapoptotic pathways (such as HER2) may emerge to drive resistance (7). Clarifying the extent of within-patient heterogeneity of ER expression could help to identify successful strategies for overcoming endocrine resistance.

Few studies have been conducted with multiple concurrent biopsies to examine lesion-level heterogeneity in ER expression. Published assessments of concurrent ER assays are generally designed for quality control or to compare techniques and tend to attribute variation to measurement or sampling error, rather than within-patient heterogeneity (8,9). Several studies have used biopsy assay to compare ER expression in the primary tumor with later recurrences. Loss of ER expression is considered to be more common than gain of ER expression (2,10,11), but...
this finding has been disputed (12,13). Comparing archived primary tumor tissue and metastatic lesions sampled at rapid autopsy, Wu et al. (14) found that patients who showed substantial downregulation of ER expression in metastases (2–10 y after primary diagnosis) tended to show downregulation in all sampled lesions. In vivo imaging using 18F-fluoroestradiol PET produces quantitative measurements of regional estrogen binding in breast cancer tumors (15). Within-patient heterogeneity in ER expression has been observed using 18F-fluoroestradiol PET (16), for 4 of 17 patients with multiple lesions scanned. 18F-fluoroestradiol PET also has potential for evaluating ER activity in bone-dominant metastatic breast cancer, a common manifestation of ER+ metastatic breast cancer that is difficult to biopsy, prone to false-negative ER by immunohistochemistry (10), and often features multiple lesions. Studies evaluating 18F-fluoroestradiol PET to predict response to ER-targeted therapy (17,18) have summarized ER expression at the patient level, as the average 18F-fluoroestradiol standardized uptake value (SUV) of target lesions. In 5 patients with 2 or more lesions analyzed, investigators at Washington University found within-patient heterogeneity in SUVs but little heterogeneity in change scores, justifying the use of averages in analysis (19). In the current study, we analyzed a larger collection of 18F-fluoroestradiol PET images to describe within-patient and between-patient heterogeneity of ER expression between tumor sites. If within-patient variation is substantial, lesion level differences in ER expression may be an important factor in endocrine therapy resistance. Characterizing the variability of ER expression between patients and within patients may also provide insight regarding the development of metastatic lesions. 18F-fluoroestradiol PET has shown promise as a predictive assay for response to endocrine therapy; this analysis addresses whether sufficient within-patient heterogeneity exists to examine response at the lesion level, as well as the patient level.

MATERIALS AND METHODS

Patient Selection
18F-fluoroestradiol PET scans for breast cancer patients scheduled to initiate, resume, or change endocrine therapy were selected for this retrospective analysis. Concomitant bisphosphonate therapy or trastuzumab was not an exclusionary criterion, but concurrent or planned cytotoxic therapy was. Additional technical requirements for inclusion were performance of a whole-body survey to evaluate all disease sites and access to an 18F-FDG PET scan conducted within 30 d of 18F-fluoroestradiol PET to assist in locating lesions. Patients were enrolled, and provided consent for, 1 of 3 18F-fluoroestradiol PET research studies. Study protocols were approved by the University of Washington Institutional Review Board and the University of Washington Radioactive Drug Research Committee.

18F-fluoroestradiol PET
18F-fluoroestradiol was prepared according to published methods (20), as reported in our prior studies (18,21). 18F was produced using either a Siemens Eclipse or a Scanditronix MC-50 cyclotron, with quality control tests performed immediately after each synthesis to evaluate radiochemical and chemical purity and to measure specific activity. Specific activity was recorded as GBq/mmol at the time of injection (mean, 184,038 GBq/mmol; range, 7,992–1,414,917 GBq/mmol). All scans were performed on an Advance PET tomograph (GE Healthcare). After an attenuation scan of 20–25 min, an average dose of 185 MBq of 18F-fluoroestradiol was injected over 2 min (range, 117–232 MBq). Sixty minutes of dynamic imaging over a single axial field of view (not used for data analysis in this study) was followed by a 5 field-of-view torso scan from neck to pelvis, with 5 min of static imaging for each axial field of view.

Lesions in 18F-fluoroestradiol PET scans were identified using paired 18F-FDG PET images, with anatomic localization and correlation (bone scan, ultrasound, MRI, or CT) for reference. Images from patients with extensive disease or ER-negative (ER−) lesions were coregistered with 18F-FDG or CT images using the PMOD (PMOD Technologies Ltd.) fusion module to accurately locate each tumor. Regions of interest (ROIs) could then be drawn on the 18F-FDG or CT image set and applied to the 18F-fluoroestradiol image set for ER− lesions, or ROIs could be drawn on the 18F-fluoroestradiol image set for extensive disease and matched to the 18F-FDG images. Separate ROIs were drawn on nonmanipulated 18F-FDG images to report 18F-FDG data. Square ROIs, approximately 1.3 cm3 (16 pixels in the ROI), were placed over the areas of highest uptake in the tumor site on 3 adjacent imaging planes, so that the total volume for each lesion was 3.8 cm3. Up to 16 lesions were identified for each patient. Lesions with diameters smaller than 1.5 cm were excluded because of partial-volume limitations. In patients with widespread bone disease, up to 3 of the largest 18F-fluoroestradiol PET lesions corresponding to the most 18F-FDG–avid lesions were tabulated for each of 3 areas: long bones, spine, and pelvis.

The SUV was calculated using the mean or maximum uptake within the ROIs defined and corrected for total weight or lean body mass (LBM):

\[
\text{SUV} = \frac{\text{tissue activity (kBq/mL)}}{(\text{injected dose/total weight})} \quad (\text{MBq/kg}).
\]

\[
\text{SUV}_{\text{LBM}} = \frac{\text{tissue activity (kBq/mL)}}{(\text{injected dose/LBM})} \quad (\text{MBq/kg}).
\]

LBM, calculated using the James method (22), was used to mitigate discrepancies between distribution volume and weight (21). The average LBM was 49 kg (range, 39–64 kg), and the average total weight was 77 kg (range, 52–127 kg). The primary analysis was conducted with LBM adjustment of SUVs, but measures using total patient weight are also reported. Similarly, the mean SUV was the primary 18F-fluoroestradiol measure, as for earlier studies, but the maximum SUV is reported as well. 18F-fluoroestradiol maximum SUV is also used for the 18F-fluoroestradiol/18F-FDG ratio because only maximum SUV measures were available for 18F-FDG PET.

Other Study Measures
18F-FDG was produced for clinical use according to standard operating procedures (23) and was supplied either by University of Washington Medical Center Radiochemistry or by PETNET Solutions. 18F-FDG PET started 45–60 min after injection of 267–402 MBq of 18F-FDG in patients who had fasted for 4 h or more. Images
were reconstructed using filtered backprojection, using the same matrix size and filter as for the 18F-fluorodeoxyglucose images. For both 18F-FDG and 18F-fluorodeoxyglucose PET scans, the maximum SUV was based on the pixel with maximum tracer uptake. Clinical information such as menopausal status, primary tumor histology, and treatment history was ascertained by chart review.

The ratio of 18F-fluorodeoxyglucose and 18F-FDG maximum SUVs was calculated to identify tumors with relatively high 18F-fluorodeoxyglucose uptake and relatively low 18F-FDG uptake, which would be susceptible to endocrine therapy (24). Conversely, lesions with a low 18F-fluorodeoxyglucose/18F-FDG ratio suggest the presence of highly proliferative, ER poorly functional disease that may be better suited to cytotoxic chemotherapy or agents targeted at growth factor pathways. The 18F-fluorodeoxyglucose/18F-FDG ratio may also control for partial-volume effects and for effects of different adjustments for distribution volume. 18F-fluorodeoxyglucose/18F-FDG ratio values are the same for weight and LBM-adjusted PET measures. A draw–redraw reader study (different from test–retest, because patients were scanned only once) was conducted to estimate the role of ROI placement in measurement error. When comparing uptake measures for 2 readers of the same lesion from the same scan, discrepancies in uptake would be due to differences in lesion identification or ROI placement within the lesion and not to instrumentation error or true site-to-site differences. Nine scans (10%) were selected at random. The primary reader provided information about the number of lesions and their anatomic location (i.e., T12) but not slice number or other scan-specific characteristics. With this information, the second reader located the lesions, drew ROIs, and recorded 18F-fluorodeoxyglucose and 18F-FDG SUVs independently of the first reader. Both readers were highly experienced in data analysis, and each had access to a different board-certified nuclear medicine physician to assist with ROI placement.

Statistical Analysis

Linear mixed-effects regression models were fitted to examine associations between independent variables (patient and lesion characteristics) and dependent variables (18F-fluorodeoxyglucose SUV, 18F-FDG SUV, and 18F-fluorodeoxyglucose/18F-FDG ratio). These models are similar to linear regression but estimate a random intercept to account for similarity between measures in the same patient, beyond the fixed-effects model (effect of covariates). SUVs were log-transformed to satisfy linear model assumptions (25), and the 18F-fluorodeoxyglucose/18F-FDG ratio underwent a square root transformation for the same purpose. The patient-level characteristics considered were the number of breast cancer lesions identified, progesterone receptor status, HER2, patient age, aromatase inhibitor therapy at the time of the 18F-fluorodeoxyglucose PET scan, and tumor histology. Each lesion site was categorized as soft-tissue/lymph node, bone, or lung. A compound symmetry covariance structure was assumed, fitting restricted maximum likelihood estimates of variance component SEs (26). The suitability of linear mixed-effects modeling assumptions was evaluated through graphical residual diagnostics and analysis of parameter estimate sensitivity to cluster-level deletion. Draw–redraw analysis covariate estimates were estimated using a linear mixed model with random effects for patient and for reader within patient. Statistical analyses were conducted using SAS/STAT (version 9.2; SAS Institute, Inc.) and R software (version 2.11.1; R Foundation for Statistical Computing).

RESULTS

Of 391 scans of 283 patients who participated in 18F-fluorodeoxyglucose PET studies between 1996 and 2006, 93 scans for 93 separate patients met the study criteria. Two scans were not available for lesion-level reanalysis because of technical difficulties. Table 1 summarizes patient characteristics for the remaining 91. One patient had a history of ER– disease but had an ER+ biopsy of metastatic disease. Twelve patients had HER2-positive disease; 6 were on trastuzumab when scanned using 18F-fluorodeoxyglucose PET. More than half of patients were known to be undergoing aromatase inhibitor therapy at the time of the scan, underscoring that many of these patients were starting or changing endocrine therapy as nth line salvage therapy. Most patients had bone lesions (67/91, 74%). Soft tissue lesions were also prevalent (54/91, 59%), with 31 of 91 patients (34%) having both bone and soft-tissue lesions.

Table 2 describes lesion-level characteristics for the 505 lesions in 91 patients. Most lesions (378/505, 75%) were bone lesions, and the remaining lesions (127/505, 25%) were located in soft tissue, lymph node, or lung. Liver lesions and lesions directly adjacent to the liver were not detectable by 18F-fluorodeoxyglucose PET because of tracer metabolism in the liver and were therefore excluded from the analysis. For the 5 early-stage patients, 7 of 10 lesions were breast lesions, and 3 were in nodes. For advanced-stage patients, 14 of the 103 soft-tissue and lymph node lesions were located in the breast. The 18F-fluorodeoxyglucose mean SUVLBM for individual lesions ranged from 0.0

\[
\text{ICC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}
\]

A high ICC value would indicate that most variability was between patients: neither measurement error nor within-patient variability in true PET measures would likely be important in using PET as a quantitative biomarker. A lower ICC could be due to either measurement error or true variability, as would occur if metastatic lesions within a patient had different molecular characteristics. If measurement error were the cause of a low ICC, the problem could be addressed by averaging data from multiple lesions, using the Spearman–Brown equation to estimate the ICC of the averaged values:
### TABLE 1
Patient and Disease Characteristics (n = 91) as Recorded at Time of 18F-Fluoroestriadiol PET Scan

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient population</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>89 (98)</td>
</tr>
<tr>
<td>Premenopausal (women only, n = 89)</td>
<td>15 (17)</td>
</tr>
<tr>
<td><strong>Weight</strong></td>
<td></td>
</tr>
<tr>
<td>Normal (body mass index &lt; 25)</td>
<td>30 (33)</td>
</tr>
<tr>
<td>Overweight</td>
<td>32 (35)</td>
</tr>
<tr>
<td>Obese (body mass index &gt; 30)</td>
<td>29 (32)</td>
</tr>
<tr>
<td><strong>Primary tumor immunohistochemistry and histology</strong></td>
<td></td>
</tr>
<tr>
<td>ER+</td>
<td>90 (99)</td>
</tr>
<tr>
<td>Progesterone receptor–positive (n = 88)</td>
<td>74 (84)</td>
</tr>
<tr>
<td>HER2/neu–positive</td>
<td>12 (13)</td>
</tr>
<tr>
<td><strong>Histology</strong></td>
<td></td>
</tr>
<tr>
<td>Ductal</td>
<td>67 (73)</td>
</tr>
<tr>
<td>Lobular</td>
<td>18 (20)</td>
</tr>
<tr>
<td>Ductal and lobular</td>
<td>6 (7)</td>
</tr>
<tr>
<td><strong>Breast cancer and treatment history</strong></td>
<td></td>
</tr>
<tr>
<td>Advanced disease</td>
<td>86 (95)</td>
</tr>
<tr>
<td>Chemotherapy for metastatic breast cancer (n = 82)</td>
<td>21 (26)</td>
</tr>
<tr>
<td>Endocrine therapy for metastatic breast cancer (n = 82)</td>
<td>38 (46)</td>
</tr>
<tr>
<td>Radiation therapy for metastatic breast cancer (n = 82)</td>
<td>28 (34)</td>
</tr>
<tr>
<td>Aromatase inhibitor therapy at time of scan (n = 90)</td>
<td>55 (61)</td>
</tr>
<tr>
<td><strong>Tumor characteristics†</strong></td>
<td></td>
</tr>
<tr>
<td>Bone lesions present</td>
<td>67 (74)</td>
</tr>
<tr>
<td>Soft-tissue lesions present</td>
<td>54 (59)</td>
</tr>
<tr>
<td>Both bone and soft-tissue lesions present</td>
<td>31 (34)</td>
</tr>
<tr>
<td>Endocrine therapy after 18F-Fluoroestriadiol PET</td>
<td></td>
</tr>
<tr>
<td>Tamoxifen</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Aromatase inhibitor</td>
<td>60 (66)</td>
</tr>
<tr>
<td>Aromatase inhibitor and fulvestrant</td>
<td>18 (20)</td>
</tr>
<tr>
<td>Other or unknown‡</td>
<td>7 (7)</td>
</tr>
</tbody>
</table>

*Patients without advanced disease (n = 5) excluded, history unknown for n = 4.
†Lesions in 18F-Fluoroestriadiol PET analysis.
‡Aromatase inhibitor therapy at time of scan (n = 90).

Data in parentheses are percentages. Mean age (±SD) of patient population was 56 ± 11 y (median, 55 y; range, 28–79 y). Mean years since breast cancer diagnosis was 7 ± 6 y (median, 5 y; range, 0–31 y). Mean years since first cancer recurrence (advanced disease only, n = 86) was 2 ± 3 (median, 0.3 y; range, 0–17 y). Mean number of lesions was 5.6 ± 4.0 (median, 4; range, 1–16).

to 10.2, with a mean value of 1.9. 18F-FDG maximum SUV<sub>LBM</sub> also showed a broad range of values for observed lesions (0.8–16.5), as did the 18F-fluoroestriadiol/18F-FDG ratio (0.1–5.6). Two patients' 18F-FDG PET scans were from outside centers and were used to identify 13 lesions but not to provide 18F-FDG SUVs for analysis. Supplemental Figure 1 (supplemental materials are available online only at http://jnm.snmjournals.org) is an example to illustrate analysis of 18F-fluoroestriadiol and 18F-FDG PET image data. Although 18F-fluoroestriadiol and 18F-FDG uptake were not uniform across all lesions, the overall impression is that the patient had aggressive (highly glycolytic) and 18F-fluoroestriadiol–avid disease.

Figure 1 illustrates the within-patient and between-patient variation in 18F-fluoroestriadiol SUV<sub>LBM</sub>. The horizontal axis is an index for each of the 91 patients, with increasing average 18F-fluoroestriadiol SUV<sub>LBM</sub> from left to right. The vertical axis is 18F-fluoroestriadiol SUV<sub>LBM</sub> on a log scale (base 2). Between-patient variation was observed: the average 18F-fluoroestriadiol SUV<sub>LBM</sub> for each patient ranged from 0.2 (upper left corner of Fig. 1) to 6.7 (lower right corner of Fig. 1). A horizontal line shows an 18F-fluoroestriadiol mean SUV<sub>LBM</sub> of 1.0, which roughly corresponds to the 18F-fluoroestriadiol mean SUV threshold of 1.5 examined in an earlier study (18). Thirty-four of 91 patients (37%) had an average 18F-fluoroestriadiol uptake below this hypothesized threshold for response to endocrine therapy, using the average of logged values, despite the fact that all patients had ER+ primary or metastatic disease by immunohistochemistry. Some patients had uniformly high 18F-fluoroestriadiol uptake (such as index 85, Fig. 2A), whereas others had uniformly low 18F-fluoroestriadiol uptake (such as index 13, Fig. 2B). A few patients showed patterns such as index 55 and 44, with gaps between SUV<sub>LBM</sub> of lesions (index 55) or clusters of lesions (index 44). Mixed 18F-fluoroestriadiol uptake between lesions is shown in Figure 2C (index 32). Of 11 recorded bone lesions, 9 had relatively high 18F-fluoroestriadiol uptake, but lesions on the sternum and right humerus had an 18F-fluoroestriadiol mean SUV<sub>LBM</sub> of less than 0.3.

Figure 3 shows within-patient and between-patient variation in 18F-FDG SUV<sub>LBM</sub> values, with the horizontal index sorted by ascending average 18F-fluoroestriadiol SUV<sub>LBM</sub>, so that index numbers are the same as for Figure 1. As for 18F-fluoroestriadiol SUV<sub>LBM</sub>, there was considerable between-patient variation, but SUVs for multiple lesions in the same patient were generally clustered close to the mean value. Figure 4 displays the 18F-fluoroestriadiol/18F-FDG ratio (with the vertical axis on a square root scale). The index 32 lesions with the lowest 18F-fluoroestriadiol SUV<sub>LBM</sub> (Fig. 2C) also had notably low 18F-fluoroestriadiol/18F-FDG ratios (Fig. 4). These observations used tabulated data, because the bone lesions cannot be tracked between figures. In contrast to index 32, for index 44 the gap between clusters of lesions for 18F-fluoroestriadiol SUV<sub>LBM</sub> (Fig. 1) was not notable in the 18F-fluoroestriadiol/18F-FDG ratio (Fig. 4), suggesting that the ratio may have controlled for between-lesion differences in size or cellularity. Linear mixed models were used to explore predictors of variation in PET parameters. Both 18F-fluoroestriadiol and 18F-FDG uptake tended to be higher in bone than in soft tissue (Table 3, P < 0.001, Wald test). However, the 18F-fluoroestriadiol/18F-FDG ratio was not associated with tumor site (P = 0.41), suggesting that the associations of tumor site with 18F-fluoroestriadiol SUV and 18F-FDG SUV.
were artifacts related to tumor size, cellularity, and contrast to surrounding tissue. In contrast, associations with the number of lesions per patient support the 18F-fluoroestradiol avidity of high-volume, bone-dominant disease. According to a univariate linear mixed model, 18F-fluoroestradiol SUV_LBM was predicted to be 27% higher for each 5 additional lesions (P < 0.01), but the number of lesions was not associated with 18F-FDG uptake or the 18F-fluoroestradiol/18F-FDG ratio in other models (P > 0.20). Holding lesion site (bone, lung, or soft tissue) constant, progesterone receptor status, HER2 overexpression, patient age, aromatase inhibitor therapy at the time of the 18F-fluoroestradiol PET scan, and primary tumor histology did not explain variation in 18F-fluoroestradiol SUV, 18F-FDG SUV, or the 18F-fluoroestradiol/18F-FDG ratio (P > 0.20). The 5 patients with early-stage disease tended to have lesions with higher 18F-fluoroestradiol SUV_LBM (P = 0.005; average 18F-fluoroestradiol SUV_LBM for bone: 3.4 for early-stage disease, 1.3 for advanced) and higher 18F-fluoroestradiol/18F-FDG ratio (P < 0.001; average: 2.3 for early-stage disease, 0.9 for advanced) but did not differ in 18F-FDG SUV_LBM (P = 0.49). In summary, fixed-effects models suggest that lesion-level differences in 18F-fluoroestradiol and 18F-FDG SUV may be artifacts of size and cellularity. Early-stage disease tended to have higher 18F-fluoroestradiol uptake than advanced disease, as might be expected because many patients with advanced disease were endocrine-refractory.

Controlling for variability in PET parameters explained by covariates in the fixed-effects model, remaining variance is accounted for by between-patient and within-patient variance components. The linear mixed models in Table 3 controlled for lesion site, disease stage, and the number of lesions, to account for the effects described. The proportion of variance explained by between-patient factors was measured by the ICC (Table 3). The ICC was 0.60 for 18F-fluoroestradiol SUV, 0.54 for 18F-FDG SUV, and 0.65 for the 18F-fluoroestradiol/18F-FDG ratio. In each case, more than half of the variation in PET parameters was explained by differences in average uptake between different patients.

The importance of remaining, within-patient, variation depends on whether the variation reflected true unmeasured differences between lesions or measurement error. We examined variability introduced by image analysts by conducting a draw–redraw study of 38 lesions in 9 patients. For most patients and most lesions, concordance on SUVs was near-perfect: the proportion of variance explained by
factors other than variation between raters was 94%, 92%, and 90%, respectively, for 18F-fluoroestradiol SUV, 18F-FDG SUV, and the 18F-fluoroestradiol/18F-FDG ratio. When lack of agreement occurred, multiple lesions per patient tended to be discrepant (Supplemental Fig. 2). To estimate the proportion of within-patient variation attributed to variation between raters, we excluded between-patient variation from the proportion and found 29% of 18F-fluoroestradiol SUV within-patient variation attributable to raters (22% for 18F-FDG SUV, 44% for the 18F-fluoroestradiol/18F-FDG ratio).

The reliability (ICC) for 18F-fluoroestradiol measures when averaging values for 2 lesions per patient was 0.75 using the Spearman–Brown equation and the 18F-fluoroestradiol SUV adjusted for lean body mass.

$$\text{FIGURE 2. Torso survey images from 3 female patients with metastatic breast cancer from ER+/HER2-negative primary histology (patient in A had PgR-negative primary tumor histology; patients in B and C had PgR-positive). (A) (sagittal view; index 85 in Figs. 1, 3, and 4): 48-y-old with bonydominant disease with high 18F-fluoroestradiol uptake and high 18F-FDG uptake. (B) (coronal view; index 13 in Figs. 1, 3, and 4): 62-y-old with uniformly low 18F-fluoroestradiol uptake on 18F-fluoroestradiol PET. Three soft-tissue lesions (18F-fluoroestradiol mean SUV$_{BM}$ of 0.3, 0.5, and 0.7) are identified by arrows. (C) (sagittal view; index 32 in Figs. 1, 3, and 4): 52-y-old with spinal lesion visible on both 18F-fluoroestradiol PET images and 18F-FDG PET images and sternal lesion with visible 18F-FDG but negative 18F-fluoroestradiol PET. FES = fluoroestradiol.}$$

$$\text{FIGURE 3. LBM-adjusted 18F-FDG maximum SUV (y-axis, log2 scale) for 492 lesions in each of 89 patients (x-axis, sorted by average 18F-FDG SUV). FES = fluoroestradiol; LN = lymph node.}$$

$$\text{FIGURE 4. Ratio of 18F-fluoroestradiol maximum SUV and 18F-FDG maximum SUV (y-axis, square root scale) for 492 lesions in each of 89 patients (x-axis, sorted by average 18F-fluoroestradiol SUV). FES = fluoroestradiol; LN = lymph node.}$$
SUV ICC of 0.60 from Table 3. Averaging 3 lesions would raise the expected reliability to 0.82, and averages of 4 and 5 lesions would be expected to have reliability of 0.86 and 0.88, respectively. Our prior strategy of averaging values for up to 3 lesions appeared to be appropriate: averaging results for seven lesions improved the precision of estimates of 18F-fluoroestradiol uptake. Improving precision is particularly useful when using a cutoff value such as an 18F-fluoroestradiol SUV of 1.0 to identify ER+ disease and guide choice of therapy.

**DISCUSSION**

This study summarized ER expression measured by 18F-fluoroestradiol PET and tumor glucose metabolism measured by 18F-FDG PET in 505 lesions scanned in 91 breast cancer patients. The ICC was 0.60 (95% CI, 0.50–0.69) for 18F-fluoroestradiol. Uptake patterns for individual patients (Fig. 1) showed that 18F-fluoroestradiol SUVs clustered around the within-patient average in most patients, suggesting that these patients’ contemporaneous lesions had similar ER expression. Averaging up to three 18F-fluoroestradiol SUV measurements from each patient should improve precision, increasing the ICC to 0.82. The 18F-fluoroestradiol/18F-FDG ratio appeared to mitigate variability attributable to partial-volume effects and variable lesion size, providing support (in addition to the biologic rationale) for exploration of the 18F-fluoroestradiol/18F-FDG ratio as a predictive biomarker for response to endocrine therapy.

The observed ICC of 0.60 may seem modest, compared with the 0.94 measure of intratumoral heterogeneity observed by Barry et al. (9), using gene expression data. However, the cohort described in this article would be expected to have less between-patient variation (σ2b) and more within-patient variation (σ2e). All were deemed candidates for endocrine therapy by the treating clinician, restricting between-patient variation in contrast to the mix of ER+ and ER− patients in the gene expression study. Within-patient variation would be expected to be higher for measures on multiple tumors in different tissue types than for repeated sampling of a single tumor.

This analysis examined quantitative 18F-fluoroestradiol results for multiple tumor sites at 1 time point per patient. All but 1 patient had ER-expressing primary tumors at diagnosis. However, a substantial fraction of patients had uniformly low or absent 18F-fluoroestradiol uptake when scanned, typically after one or more endocrine-targeted therapy regimens. This broad between-patient variation in average 18F-fluoroestradiol uptake and 18F-fluoroestradiol/18F-FDG ratio supports the hypothesis that recurrent breast cancer may not share the phenotype of the primary tumor (10).

We found a small number of patients with highly discordant 18F-fluoroestradiol uptake across tumor sites, sug-
suggesting that this phenomenon is relatively rare. These patients appeared to have both ER+ and ER− lesions at the same time, demonstrated by discordant findings on 18F-fluoroestradiol and 18F-FDG PET. This observation suggests that low or absent ER expression may emerge relatively quickly in the course of breast cancer progression, perhaps under the selective pressure of endocrine therapy. The examination of molecular features in patients with extreme site-to-site heterogeneity, with tissue sampling directed by imaging, may therefore yield important clues as to the nature of the emergence of breast cancer with low ER expression. Future work will include analysis of the disease history, treatment history, and prognosis of patients with mixed 18F-fluoroestradiol uptake, such as the patient in Figure 2C.

A limitation of this study is the lack of concurrent biopsy tissue, to compare 18F-fluoroestradiol PET results with immunohistochemistry or gene expression assessment of ER. However, the impracticality of serial and multiple biopsies in metastatic breast cancer patients is one of the rationales for this study, and other work has demonstrated the association between 18F-fluoroestradiol SUV and ER by immunohistochemistry (15). 18F-fluoroestradiol PET was performed only at 1 time point for each patient, so within-patient changes in ER expression over time could not be addressed. Another limitation is that the PET studies were conducted on a heterogeneous group of patients, with different treatment histories and prognoses. Prospective studies for using 18F-fluoroestradiol PET to guide endocrine therapy are planned, and this retrospective evaluation helps to resolve design questions (i.e., how many lesions to assess) in the conduct of those studies.

CONCLUSION

Currently, selection of patients for endocrine therapy is based on ER expression as determined from biopsy material, often from a temporally distant primary tumor. Improved identification of ER expression in advanced breast cancer can better select patients who are likely to benefit from hormone therapy and may identify targets for the development of future therapies. Imaging ER function by PET can be exploited to identify, characterize, and follow treatment response for multiple lesions in the same patient. 18F-fluoroestradiol PET is particularly promising in the metastatic setting, in which patients have many lesions in areas (such as bone) that are difficult to biopsy. Future work will examine patients with highly discordant 18F-fluoroestradiol uptake across tumor sites and will evaluate the role of between-patient and within-patient variability in 18F-fluoroestradiol SUV and the 18F-fluoroestradiol/18F-FDG ratio in predicting clinical response to endocrine therapy.

DISCLOSURE STATEMENT

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

ACKNOWLEDGMENTS

We thank Svetlana Stekhova, Gina Massimino, Xiaoyu Chai, Mark Muzi, the study participants, the referring physicians, and the nuclear medicine team in the Department of Radiology. We gratefully acknowledge financial support from a Breast Cancer Research Program Pilot grant (Safeway) and NIH grants CA42045, CA72064, CA148131, CA138293, and DOD W81XWH-04-01-0675. No other potential conflict of interest relevant to this article was reported.

REFERENCES


